

Clase auxiliar # 2

Software Stata, teorema de Frisch-Waugh e interpretación

¡Recordar!

Matriz de proyección: $P_X = X(X^T X)^{-1} X^T$

Matriz de aniquiladora o de proyección ortogonal: $M_X = I - P_X$

Teorema de Frisch-Waugh: en los dos modelos que se enuncian a continuación, la estimación por MCO de β_B es equivalente. Es decir, $\hat{\beta}_B^{\text{Modelo 1}} = \hat{\beta}_B^{\text{Modelo 2}}$.

- Modelo 1: $Y = \vec{\beta}_0 + X_A \beta_A + X_B \beta_B + U$
- Modelo 2: $\underbrace{M_{X_A} Y}_{\hat{U}_{Y, X_A}} = \underbrace{M_{X_A} \vec{\beta}_0}_{\hat{\beta}_0'} + \underbrace{M_{X_A} X_B}_{\hat{U}_{X_B, X_A}} \beta_B + \underbrace{M_{X_A} U}_{U'}$

$\hat{U}_{W, Z}$ es el residuo obtenido al regresionar W en función de Z .

Juguemos con Stata

Ocuparemos la base de datos Casen 2015.dta y el código Auxiliar 02 - Gasto en educación.do (ver Material Docente). Antes que todo, presentaremos nuestras variables de interés:

- `expr`: se interpreta como la cantidad de personas en la población, que representa una persona en la muestra. Sirve para darle más peso a unas observaciones que a otras.
- `folio`: identificador único del hogar. En la base de datos, cada observación es una persona, por lo que distintas observaciones pueden tener un mismo folio.
- `edad`: edad de la persona.
- `e13b`: gasto mensual en educación parvularia, básica o media.
- `e15a`: gasto mensual en educación técnica superior, profesional o posgrado.
- `qaut`: quintil al que pertenece el hogar.
- `pco1`: parentesco con el/la jefe(a) de hogar (`pco1 = 1` \Rightarrow la persona es el/la jefe(a) de hogar).
- `ytoth`: ingreso total del hogar.

1. **Tratamiento de la base de datos:** mediante el comando `keep` deje sólo las variables de interés. Usando `gen` y `replce` cree variables que almacenen el gasto total en educación, la edad del jefe de hogar y el cuadrado de la edad del jefe de hogar. Finalmente, usando `collapse` haga que el nivel de agregación de la base sea por hogar y no por persona.
2. **Regresiones**¹: escriba matemáticamente las siguientes regresiones y ejecútelas en *Stata* considerando sólo hogares donde el/la jefe(a) tenga entre 25 y 65 años.

Regresión 1: gasto en educación del hogar en función de la edad del jefe de hogar y el cuadrado de la edad del jefe de hogar.

Regresión 2: gasto en educación del hogar en función del ingreso total del hogar, la edad del jefe de hogar y el cuadrado de la edad del jefe de hogar.

Regresión 3: gasto en educación del hogar en función del ingreso total del hogar, la edad del jefe de hogar y el cuadrado de la edad del jefe de hogar, todo esto clasificado por quintiles.

A partir de ahora, sólo considere los resultados de la **segunda regresión**.

3. **Interpretación de resultados:** comente resultados relacionados a los signos de los coeficientes, significancia individual, significancia global, número de observaciones y ajuste del modelo.
4. **Gráfico:** bosqueje el modelo mediante un gráfico. Notar que se trata de una función de \mathbb{R} a \mathbb{R}^2 .
5. **Retornos marginales esperados:** obtenga el retorno marginal esperado sobre el gasto en educación causado por el ingreso del hogar y por la edad del jefe(a) de hogar cuando éste(a) tiene 30 y 40 años respectivamente.
6. **Teorema de Frisch-Waugh:** usando este teorema, obtenga $\hat{\beta}_{\text{ingreso}}$.
7. **Modelo log-lin:** ahora realice una regresión análoga a la segunda, donde la variable dependiente sería esta vez el logaritmo del gasto en educación. Interprete el coeficiente β_{Ingreso} y justifique su interpretación.

¹Podríamos cuestionarnos la causalidad que tienen las variables consideradas, sin embargo por esta vez la asumiremos.